Remote Data Integrity Protection Enhancement in Cloud Storage

Varsha B.
Department of M.Tech CSE
Dr. Ambedkar Institute of Technology, Bangalore.

Abstract— Cloud computing is the delivery of computing services over the Internet. Protecting all users' data in the cloud is highly challenging. Since all our data stored in third party's computer, security concerns arise. The main focus in cloud storage is data security and integrity. This paper works on securing cloud data from corruptions, checking for its integrity and recovering data during failures and corruptions.

Keywords—Cloud Storage, Security, Integrity checking, Failure Recovery

I. INTRODUCTION

Cloud is the set of hardware, software, networks, storage, services and interfaces that combine to deliver aspects of computing as a service and Cloud computing is delivering these services to end-users whenever and wherever they need it over the Web.

Cloud storage is pool where all data and applications are stored which are hosted by Cloud Service Provider. Cloud storage has gained popularity due to its low storage price and high accessibility. Cloud storage key characteristics are storage service delivered over a network (internet or intranet), Easy to scale, Easy to manage. In spite of many benefits, cloud faces problem like data integrity, loss etc.

Data integrity is a major cloud storage issue. Integrity is the guarantee that the data is correct, accurate and unchanged. Checking the integrity in cloud storage leads to problems like transmission consumes large bandwidth, higher IO cost due to huge data volume of user, usage of small devices (PDA, Cell phone) by data owner.

Storing all data in a single server leads to single-point-of-failure problem and vendor lock-ins. So the data is stripped onto the multiple servers. This stripping method is highly advantageous (e.g. Read data from surviving servers i.e. repair traffic, reconstruct the corrupted data of the failed server and writing the reconstructed data on new server) while repairing the failed nodes or servers.

II. LITERATURE REVIEW

Data storage in cloud can be modified by anyone. While storing data at cloud storage space we are restricted by the resources at cloud storage servers [3]. To secure the cloud storage data, the owner has to apply cryptographic techniques and perform frequent checks. Proof of retrievability (POR) and Proof of data possession (PDP) are proposed to verify the integrity of large files [1]. The parity scheme in RAID is used for recovering from a single disk failure in the array. Error correction codes are used to detect and correct bit errors in storage devices [5].

III. PROBLEM STATEMENT

The main idea is to avoid client retrieving complete file from the server. If the corruptions are detected (e.g. when servers are down or when the data is modified maliciously) then the corrupted data has to be repaired using remaining servers' data and replace the original data in the new server.

IV. PROPOSED SYSTEM

Consider cloud setup, a single proxy server connected to six storage servers shown in Fig-1. The proxy node is responsible for coordinating with the storage servers and replying to the client with appropriate messages. The storage servers provide on-disk storage. Basic operations like encode, decode, check, repair and delete has to be performed by the user from the proxy node on the storage nodes.

Initially the authorized user identifies him with the identity credentials. This enables user to access the storage systems and perform the suitable basic operations. When the user encodes the files using AES-128 encryption method, the files are distributed onto the storage nodes as file chunks (striping).The decoded files can be stored in the temporary storage server (e.g. node n6). The user performs the check operation by using the parities and applies error correction. Then the corrected chunk is verified with its Message Authentication Code. The user can delete the encoded files from all the nodes whenever necessary. Finally the user has to

ISSN: 2278 – 1323

repair the failed node or corrupted node in the spare nodes (e.g. Consider chunks are distributed on nodes n1, n2, n3and n4, if one server fails then it has to repair the corrupted data and store it onto n5. Now the node n5 behaves as a storage node that failed.) Thus the fault tolerance is achieved and failure recovery is performed. Notifying user about the intentional or unintentional modification during the access time increases the more secure and trustworthy environment.

.

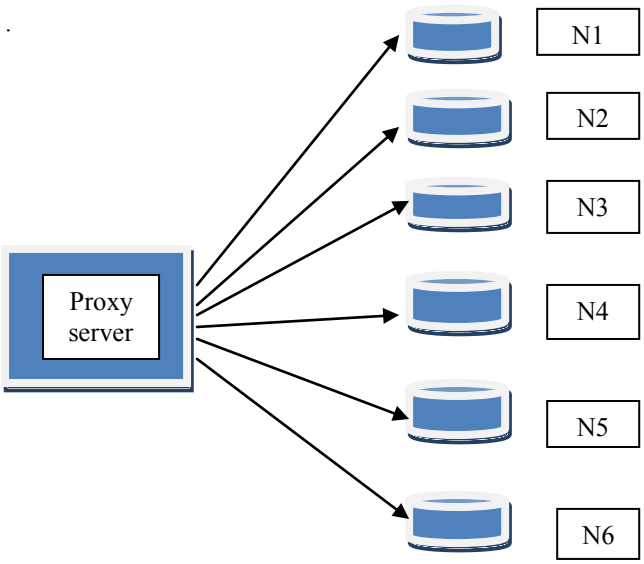
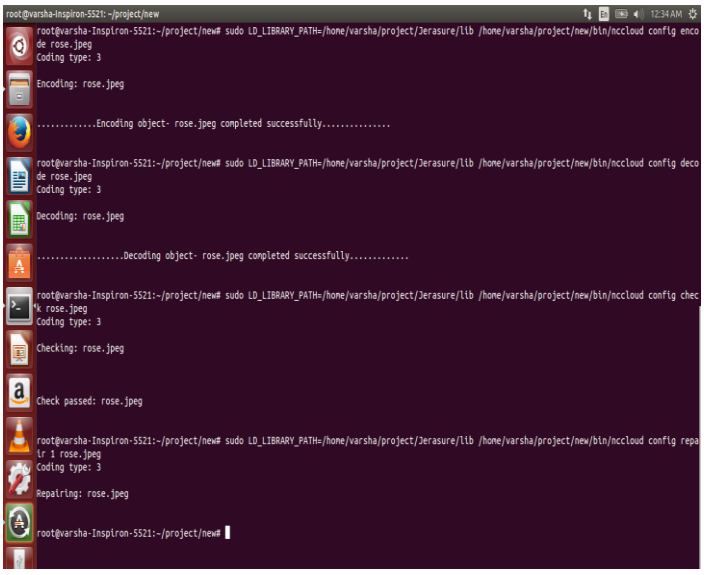
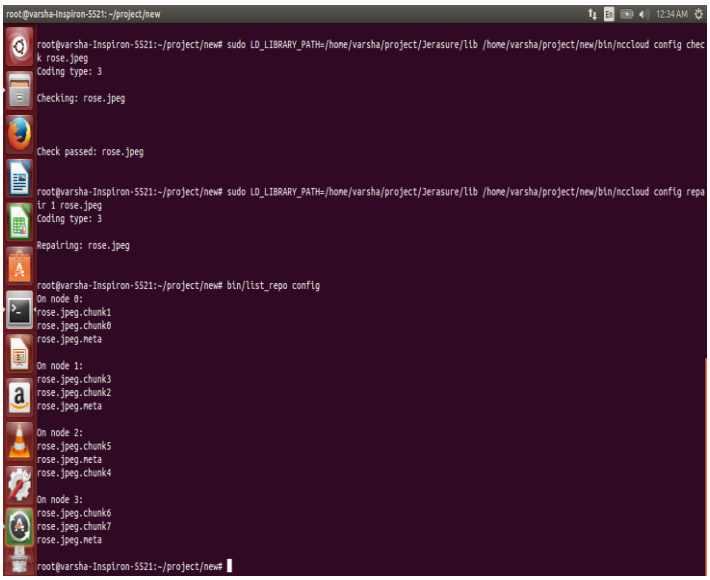


Fig-1: Single proxy server connected to six storage servers which are labeled n1 to n6 respectively.

The experiments were conducted on cloud platform built on Openstack. The screenshots attached below shows the expected results of basic operations like encode, decode, delete, check and repair.





V. CONCLUSION

Storing all data in cloud is not trustworthy. So in order to protect data in cloud, the data has to be encrypted and its integrity check has to be performed. During the times of failures and corruption the data has to repaired and restored.

REFERENCES

[1] A Survey on Data Integrity on Cloud Storage in Cloud Computing. JAFRC, Feb 2014
[2] Identifying Data Integrity in Cloud Storage. IJCSI, March 2012
[3] Data Integrity Proofs in Cloud Storage. JCT 2014
[4] NCCLOUD: Applying Network Coding for the Storage Repair in a Cloud-of-Clouds. IEEE, 2012
[5] Ensuring Data Integrity in Storage: Techniques and Applications. Gopalan Sivathanu, Charles P. Wright, and Erez Zadok Stony Brook University.
[6] Storage Defined Storage with Openstack Swift by Joe Arnold
[7] Cloud Computing. Dr. Kumar Saurabh

ISSN: 2278 – 1323